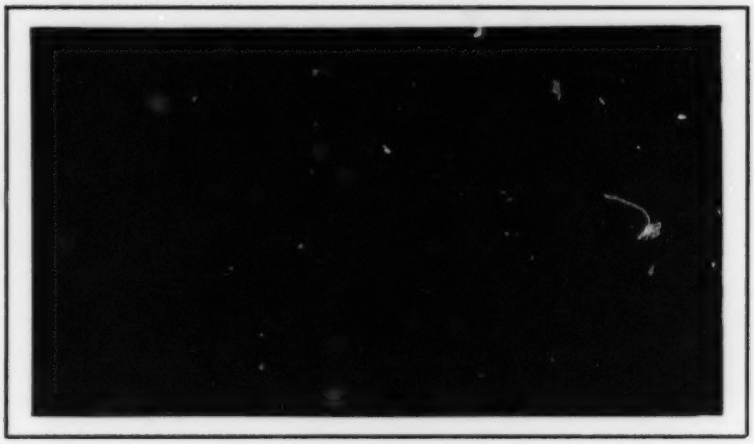

CENTRE FOR
SYSTEMS
SCIENCE

Simon Fraser University, Burnaby, B.C., Canada V5A 1S6 / (604) 291-4704

92-05402/

Computing The Lower Bounded Composite Hypothesis By Belief Updating

Yang Xiang

CSS-IS TR 92-01

January 1992

Computing The Lower Bounded Composite Hypothesis By Belief Updating

Yang Xiang

Expert Systems Laboratory
Centre for Systems Science
Simon Fraser University
Burnaby, B.C., Canada, V5A 1S6
yangx@cs.sfu.ca

Abstract

This paper presents a threshold method of computing Lower Bounded Composite Hypotheses (LBCHs) by belief updating (computation of Posterior Marginal Distributions (PMDs)) in Bayesian networks. The LBCHs are required when decisions are to be made with reward/penalty type utilities. LBCHs are Most Likely Composite Hypotheses (MLCHs) if the lower bound is greater than 0.5. Although the method is practical only for computing LBCHs over small number of (about 20) variables (with no restriction to the size of the domain), it is argued that this can often be justified and can even be desirable. There has been no known algorithm computing MLCHs in clique tree based secondary structures of Bayesian networks. The method presented allows the computation of LBCHs in such structures using existing belief updating algorithms. An example is given to illustrate the method.

1 Introduction

Bayesian belief networks combine graphic representation of causal domain models and probability theory. Over the last decade, they have gained increasing popularity as a natural, efficient knowledge representation method and a consistent inference formalism for building knowledge based systems which require reasoning under uncertainty.

Bayesian networks have been termed in the literature as *belief networks*, *causal networks*, or *causal probabilistic networks*. Formally a Bayesian network [Pearl 88] is a triplet (N, E, P).

- The *domain* N is a set of nodes each of which is labeled with a random variable. Each variable has a set of mutually exclusive and exhaustive outcomes which forms its space. 'Node' and 'variable' are used interchangeably in the context of Bayesian nets. A joint assignment of outcomes for a set of variables $X = \{A_1, \ldots, A_n\} \subseteq N$ is denoted by the concatenation of corresponding outcomes $x = a_1 \ldots a_n$.

- E is a set of arcs such that (N, E) is a directed acyclic graph. The arcs signify the existence of direct causal influences between the linked variables. Causality is directed along the directed arcs. The basic dependency assumption embedded in Bayesian nets is that a variable is independent of its non-descendants given its parents.

- P is a joint probability distribution quantifying the strengths of the causal influences signified by the arcs. P specifies for each $A_i \in N$ the distribution of the random variable labeled at A_i conditioned by the values of A_i's parents π_i in the form of a conditional probability table $p(A_i|\pi_i)$. $p(A_i|\pi_i)$ is a normalized function mapping the joint space of $\{A_i\} \cup \pi_i$ to $[0, 1]$. The joint probability distribution P can be represented as $P = p(A_1, \ldots, A_n) = \prod_{i=1}^{n} p(A_i|\pi_i)$.

Inference in Bayesian networks can provide answers to essentially two types of queries. The first type asks the PMDs in the form: $p(A|\mathbf{e})$ where A is a *single* variable and $\mathbf{e}$ stands for all the evidence acquired so far. Computing PMDs is termed *belief updating* [Pearl 88]. A PMD allows one to rank the set of possible outcomes for variable A according to their credibility given $\mathbf{e}$, for example, whether a patient has a particular disease or not; or whether a component in a machine is faulty.

The second type asks the MLCH. Let $X \subseteq N$ be a set of variables. One computes a joint assignment x of outcomes to all variables in X such that

$$(\forall x')p(x|\mathbf{e}) > p(x'|\mathbf{e}).$$

A joint assignment x' is called an *instantiation* of X, or a *composite hypothesis*, or a *configuration*, or an *explanation*. Computing the MLCH is termed *belief revision* [Pearl 88]. The MLCH is needed when one tries to determine the exact state of the world (when $X = N$) [Poole and Provan 90], for example, to give a description of the overall status of a patient.

In the literature, the two types of inference have been considered distinct. Consider two variables A and B. The computation for PMDs may indicate that a_1 is the most likely outcome of A and b_1 is the most likely outcome of B. While the MLCH for $X = \{A, B\}$ may turn out to be $a_1 b_2$ [Pearl 88, Poole and Provan 90].
This paper shows when this is not the case. Based on the distinct nature of the two inference types, separate algorithms have been developed for each type of inference. For computation of *exact* PMDs, there are $\lambda - \pi$ message passing in trees [Pearl 86], arc-reversal and node-removal [Shachter 86, Shachter 88], directed clique tree message passing [Lauritzen and Spiegelhalter 88], junction trees of cliques [Jensen, Lauritzen and Olesen 90a, Jensen, Olesen and Andersen 90b], multiply sectioned Bayesian networks and junction forests of cliques [Xiang, Poole and Beddoes 91]. For computation of *approximate* PMDs, there are logic sampling [Henrion 88], and clamped logical sampling [Pearl 88]. For computation of the MLCH,

there are message passing, cutset conditioning, clustering [Pearl 88] as *exact* methods; and TopN [Henrion 91] as an *approximate* method.

This paper studies the relation between the two types of inference. In particular, the paper presents the following. (1) The notion of *Lower bounded Composite Hypothesis* (LBCH) is introduced from a decision theoretic context. LBCHs are MLCHs when the lower bound is greater than 0.5. (2) It is shown that LBCHs can be obtained by belief updating. (3) Although computation of LBCHs by belief updating is directly applicable to hypothesis sets of less than about 20 variables, it is argued that not only this is often sufficient with respect to one's current attention, but this is desirable compared to computing a MLCH over all domain variables. (4) There is no known algorithm computing MLCHs in clique tree based secondary structures of Bayesian networks. By establishing the relation between belief updating and revision, one can extract extra information (LBCHs) beyond PMDs from the secondary structures in a straightforward way.

Section 2 introduces the LBCH notion from maximum expected utility principle of decision theory. Section 3 and 4 presents the threshold method for computing LBCHs by belief updating. Section 5 discusses the application of the method for inference in secondary structures. Section 6 demonstrates the method by an example. Section 7 summarizes the main results of the paper.

2 Lower Bounded Composite Hypotheses

Decision theory argues a rational decision maker should act according to Maximum Expected Utility (MEU) principle [von Neumann and Morgenstern 47]. Following the formulation in [Andreassen et al. 89], the decision maker's preference can be represented by a utility function $U(x_i|x_j)$ where x_i and x_j are both composite hypothesis on a set X of variables. The function specifies the utility of action based on x_i given that x_j is true. We consider reward/penalty type of utility functions. The utility of a correct decision ($i = j$) is positive (reward). The utility of at least one wrong decision ($i \neq j$) is negative (penalty). Utility functions in many applications belong to this type. Military actions on battlefields are such examples. In medical diagnoses, treatments (medication, surgery) based on wrong diagnoses can cause side effects and inappropriate follow-ups, can prolong and worsen patients' suffering due to untreated real disease. There has been argument [Huchlenbroich 91] that medical expert systems should base their conclusion on "conclusive" evidence to avoid incorrect diagnosis as much as possible.

The MEU principle says that given evidence e, and utility function U, one should choose the action indicated by x_i which has maximal expected utility:

$$\overline{U}(x_i) = \max_k (\sum_j U(x_k|x_j)p(x_j|\mathbf{e}))$$

Denote the reward/penalty type utility as the following.

$$U(x_i|x_j) = \begin{cases} V & i = j \\ w_j & i \neq j \end{cases}$$

where $\min(w_j) = -W$, and V, W are positive. The expected utility for x_i is

$$\begin{aligned} \overline{U}(x_i) &= Vp(x_i|\mathbf{e}) + \sum_{j \neq i} w_j p(x_j|\mathbf{e}) \\ &\geq Vp(x_i|\mathbf{e}) - W(1 - p(x_i|\mathbf{e})) \\ &= W(1+r)(p(x_i|\mathbf{e}) - p') \end{aligned}$$

where $r = V/W$ (*reward/penalty ratio*) and $p' = 1/(1+r)$.

Definition 1 *With reward/penalty type utility, a composite hypothesis x is* **profitable** *if and only if $\overline{U}(x) > 0$.*

If $p(x_i|\mathbf{e}) > p'$, then x_i is profitable.

Consider two extreme cases of the reward/penalty ratio. If $V << W$, then $p' \to 1$. In this case, trying to find the profitable composite hypothesis means trying to avoid heavy penalty. For instance, when a medication can cure a patient if he has the disease but can kill him if the diagnosis is wrong, the physician should be very careful to ascertain the diagnosis. On the other extreme, if $V >> W$, then $p' \to 0$ and the right side of the above equation is approximately $Vp(x_i|\mathbf{e})$. In such case, many composite hypotheses are profitable but the reward associated may be very low compared to V. If one expects minimum reward V_0, then one should try to find x_i such that $p(x_i|\mathbf{e}) > V_0/V$. To summarize, with reward/penalty type utilities, one has a utility dependent lower bound

p_0 such that one needs to find composite hypothsi(e)s x_i satisfying $p(x_i|\mathbf{e}) > p_0$. In the about two extreme cases, p_0 takes the values of p' and V_0/V respectively. We call such x_i *Lower Bounded Composite Hypothesis* (LBCH).

LBCHs may not be unique if $p_0 < 0.5$.

Definition 2 *A composite hypothesis x is* **dominant** *if and only if $p(x|\mathbf{e}) > 0.5$.*

Proposition 1 *A LBCH is a MLCH and is unique if it is dominant.*

Since dominant LBCHs are profitable, unique and most likely, they are more useful than non-dominant LBCHs.

Given a state of knowledge, there may not exist any LBCH relative to a particular p_0. One can, however, strive for LBCHs by collecting more evidence to increase the amplitude of $p(x_i|\mathbf{e})$.[1]

3 Computing LBCHs with Belief Updating

This section derives the condition under which belief updating can be used to compute LBCHs.

Consider a set $X = A_1 \dots A_n$ of binary (this restriction will be eliminated latter) variables and an instantiation x of X. The following lemma provides a sufficient condition for computing PCHs.

Lemma 1 *Let $x = a_1 \dots a_n$ be an instantiation of a set X of n binary variables. x is a LBCH relative to p_0 if $p(a_i|\mathbf{e}) > (n - 1 + p_0)/n$ ($i = 1 \dots n$).*

Proof:

To simplify the notation, write $p(a_1 \dots a_n|\mathbf{e})$ as $p_{1\dots n}$, and $p(\overline{a_1}a_2 \dots a_n|\mathbf{e})$ as $p_{\overline{1}\dots n}$ where the bar means 'not'. Let $T = (n - 1 + p_0)/n$.

$$
\begin{aligned}
\sum_{i=1}^{n} p_i \;=\; & (p_{1\dots n} + \underbrace{p_{\overline{1}2\dots n} + \dots + p_{12\dots\overline{n}}}_{C_{n-1}^1} + \\
& \underbrace{p_{1\overline{2}3\dots n} + \dots + p_{1\dots\overline{n-1}n}}_{C_{n-1}^2} + \dots + \underbrace{p_{1\overline{2}\dots\overline{n}}}_{C_{n-1}^{n-1}} \\
& + \dots + \\
& (p_{1\dots n} + \underbrace{p_{\overline{1}\dots n} + \dots + p_{1\dots\overline{n-1}n}}_{C_{n-1}^1} + \\
& \dots + \underbrace{p_{\overline{2}\dots\overline{n-1}n}}_{C_{n-1}^{n-1}}) \\
= \; & n(p_{1\dots n}) + \\
& (n-1)(\underbrace{p_{\overline{1}\dots n} + \dots + p_{1\dots\overline{n}}}_{C_n^1}) + \\
& (n-2)(\underbrace{p_{\overline{12}\dots n} + \dots}_{C_n^2}) + \\
& \dots + \\
& 1(\underbrace{p_{\overline{1}\dots\overline{n-1}n} + \dots}_{C_n^{n-1}}) \\
> \; & nT
\end{aligned}
$$

Since the terms in the parentheses in the second equation are all different, the above implies

$$p_{1\dots n} > nT - (n-1) = p_0$$

□

A multiple value variable A can be transformed into a binary variable A', with the outcome a of A in which one is interested as one outcome of A', and with other outcomes of A congregated as the other outcome of A'. Since such transformation is always possible, Lemma 1 can be generalized to multiple value variables.

Lemma 1 establishes a threshold for each $p(a_i|\mathbf{e})$. Once each variable in X has one outcome passing the threshold, the conjunction of this set of outcomes forms a LBCH on X. It can be seen from the above proof that the threshold for each outcome is not necessary. Replacing such threshold by a threshold for the sum of $p(a_i|\mathbf{e})$, one has the slightly stronger result ((2) in the following theorem).

Theorem 1 *Let $x = a_1 \dots a_n$ be an instantiation of a set X of n variables. x is a LBCH relative to p_0 if*

1. *(individual threshold) $p(a_i|\mathbf{e}) > (n - 1 + p_0)/n$ ($i = 1 \dots n$) or*
2. *(sum threshold) $\sum_{i=1}^{n} p(a_i|\mathbf{e}) > n - 1 + p_0$.*

The following table shows the threshold for (1) in theorem 1 in relation with n when $p_0 = 0.5$.

4 How Restrictive is the Threshold Method?

Table 1 shows that computation of LBCH (with $p_0 = 0.5$) of more than about 20 variables by

[1] For simplicity of presentation, we assume the cost of collecting evidence is much smaller than $V + W$.

n	2	3	5	10	20
(n-0.5)/n	0.75	0.83	0.90	0.95	0.975

Table 1: Threshold for marginal probabilities

belief updating requires above 0.975 posterior probability for each outcome involved. This seems to be too tight a restriction on the size of the application domain. Fortunately, although the number of variables of an application domain can easily exceed 20, computation of LBCHs over less than about 20 variables can usually be justified and can even be desirable. Since dominant LBCHs are also MLCHs (Proposition 1) and are more useful than non-dominant LBCHs, the following discussion refers majorly to MLCHs to match the concept used in the referenced literature. We specify the conclusion applicable to LBCHs whenever possible.

4.1 Belief revision over all variables is not desirable

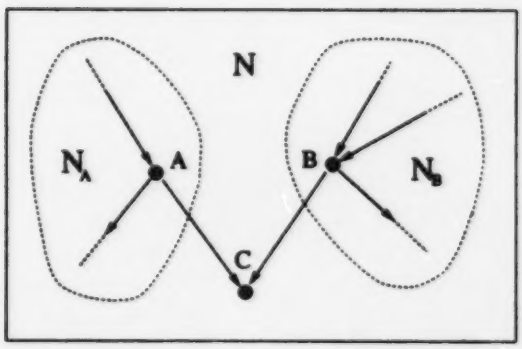

Figure 1: Illustration of problems with belief revision

The difficulties associated with belief revision over all variables have been identified in [Pearl 88, Poole and Provan 90]. First of all, there is the computational problem. If the MLCH is to cover the entire domain, every piece of evidence must be propagated to the entire network. Unobserved consequence variables no longer block the propagation as is the case in belief updating. For example (Figure 1), if variable C is unconfirmed, the subnetwork N_A is independent of N_B if PMDs are to be computed. However, in the case of belief revision over all variables, when evidence is available in N_A, it has to be propagated to N_B in order to find the MLCH. This would impose heavy computational overhead in larger applications.

Secondly, there is the conceptual problem. In belief revision, changing the space where MLCH is computed will change the outcome for each variable. For example (Figure 1), assuming all evidence is contained in N_A, and A, C are uncertain, computing the MLCH in N_A or in N may end up with different assignment for variable A. When a diagnosis problem is coupled with holiday planning, imagination of the uncertainty of trips can render a highly likely normal diagnosis to become positive (with disease) as pointed out by Pearl [1988].

The above problems suggest that computation of the MLCH over all variables of the domain is not desirable.

4.2 Variables which should be covered by a MLCH

In order to overcome the problems associated with belief revision over all variables, Pearl [1988] proposed to circumscribe a domain into a set of variables called *explanation corpus*, and to compute the MLCH only over the corpus. According to Pearl, the corpus includes evidence and ancestors of evidence variables which have a significant impact on pending decisions. Restricting computation of the MLCH on the corpus will eliminate a substantial number of variables in the domain. In this subsection, we take a closer look at the variables included in the corpus to further focus the computation to an even smaller set of variables.

Typically a known evidence variable takes one of its outcomes. Therefore, within the corpus, one can further eliminate the evidence variables from computation of the MLCH. The remaining variables are the ancestors of the evidence variables. These variables can be organized in different levels of abstractions. For example, in medical applications, a Bayesian net can have symptom variables as leaves. The successive next higher levels can be biological state level, disease or syndrome level, and disease groups level (representing diseases with some common characteristics). Not all of them are of a physician's current interest. Usually, only variables at certain level of abstraction is of physician's utmost interest. The number of such variables

will be much smaller than the number of variables within the explanation corpus.

When the number of variables in the abstraction level of current attention is large, an important strategy can be applied as suggested in [Patil, Szolovits and Schwartz 82]. One can move up an abstraction level where the important differences are represented by a smaller set of higher level hypotheses. After gathering further evidence such that the differentiation at the higher level is resolved, one can move down to the lower level to differentiate in greater detail. Focusing one's attention in this fashion, one needs only to compute the MLCH on a small number of variables.

The existence of a LBCH over a set of variables is relative to a adequate amount of evidence. Thus it is often useful to know there is no LBCH over a set of variables given the current state of knowledge. The following theorem gives one condition. Its proof is trivial.

Theorem 2 *Let X be a set of n variables. Let $A \in X$ have m possible outcomes $\{a_1, \ldots, a_m\}$. If $p(a_i|e) < p_0$ $(i = 1 \ldots m)$, then none of the instantiations of X is a LBCH relative to p_0.*

Theorem 2 asserts that there is no LBCH if A is involved in the composite hypothesis. That is, given the available evidence, one can not make any strong commitment on any composite hypothesis which involves A. Perhaps, one could choose to collect more evidence in order to sharpen the distribution and to make a better informed decision or to make a decision with A ignored. Both can be seen in medical diagnosis. When a doctor does not have adequate evidence to give a positive diagnosis on a disease, he/she may either perform more tests on the patient or give no treatment until the disease further manifests.

Computationally, one specifies the abstraction level of current interest in the explanation corpus. To compute dominant the LBCH, one uses Theorem 2 to exclude the variables at the current level which do not lead to a LBCH. Suppose a set X of M variables remains after the exclusion. Using the threshold in theorem 1, one can find (with trivial computation) all dominant LBCHs covering a subset of 2 variables in X, 3 variables, ..., up to $n \leq M$ variables, where n increases as more evidence is gathered. The computation for non-dominant LBCHs is similar.

5 Computation of Dominant MLCH in Secondary Structures

Inference in Bayesian nets can be performed in original networks [Pearl 86, Shachter 88, Henrion 88, Pearl 88, Henrion 91] or in clique tree based secondary structures [Lauritzen and Spiegelhalter 88, Jensen, Lauritzen and Olesen 90a, Xiang, Poole and Beddoes 91]. The above threshold method for computing LBCHs can be used in conjunction with algorithms of belief updating in both structures. Evidential reasoning in secondary structures have advantages when the domain knowledge are acquired once but used for multiple cases (e.g. expert systems); and when the net topology is multiply connected. To the author's knowledge, however, there has been no general algorithm which computes MLCHs in clique tree based secondary structures. All known algorithms for secondary structures perform belief updating and compute PMDs. The threshold method of Theorem 1 allows one to obtain the LBCHs in secondary structures using existing belief updating algorithms.

An alternative for computing MLCHs (or dominant LBCHs) using belief updating is to introduce auxiliary nodes [Pearl 88]. For a set X of variables, an auxiliary node A with all the possible joint assignments xs as its outcomes is added to the original net. The method compute both the MLCH and its probability. However, it is difficult to integrate with clique tree based secondary structures. For each new X and corresponding auxiliary node, a recompilation of the original net into a secondary structure is required. Since an auxiliary node introduces possibly several new loops, its addition can also complicate the secondary structure and increase the computational complexity.

Compared with the auxiliary node method, the threshold method computes only LBCHs but not their probabilities. But it does not require any recompilation of the secondary structure and supports dynamic change of the set of variables on which LBCHs are to be computed. The extra computation beyond belief updating is trivial.

6 An Example

This section provides a simple circuit diagno-

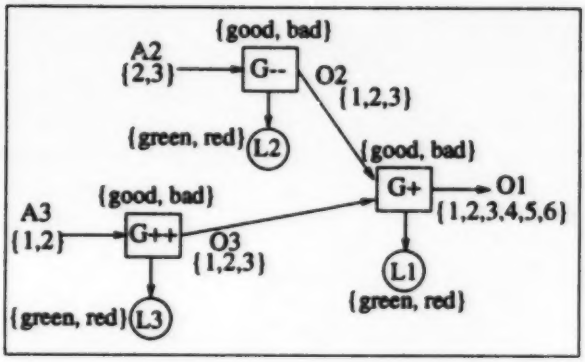

Figure 2: three gates circuit example

sis example to illustrate the threshold method.

The circuit consists of three gates: G−−, G++, and G+ (Figure 2). G−− decreases its input A_2 by one. G++ increases its input A_3 by one. G+ adds the output O_2, O_3 of G−− and G++ to produce O_1. The domains of input and output of each gate is illustrated in Figure 2. Each gate has a prior probability of being faulty (0.03, 0.02 and 0.01 for G+, G−−, and G++ respectively). When G−− and G++ are faulty, their output is clamped to the input with 98% of chance; and can be any other values in their output domain with 1% of chance for each value. When G+ is faulty, its output is clamped to the input of greater value with 95% of chance; and can be any other values in its output domain with 1% chance for each value. Suppose one has no access for O_2, O_3, but can observe an indicator light for each gate (L_1 for G+, L_2 for G−−, and L_3 for G++). Each light monitors the status of the corresponding gate with 'red' indicating a faulty state, and 'green' a normal state. The lights are not totally reliable though. When the gate is faulty, only 70% of time, the light turns red. When the gate is normal, there is still 10% of chance the light gives false alarm (red).

The circuit is represented by a Bayesian net in Figure 3. The probabilistic assumptions are summarized below. Histograms show the PMDs with the bottom level as 0 and top level as 1. Figure 3 to 6 are produced using the ***WEBWEAVR*** expert system shell [Xiang et al. 91].

$$
\begin{aligned}
p(G+ = bad) &= 0.03 \\
p(G-- = bad) &= 0.02 \\
p(G++ = bad) &= 0.01 \\
p(A2 = 2) &= 0.5 \\
p(A3 = 1) &= 0.5 \\
p(O2 = A2 - 1 | A2, G-- = good) &= 1 \\
p(O2 = A2 | A2, G-- = bad) &= 0.98 \\
p(O2 \neq A2 | A2, G-- = bad) &= 0.01 \\
p(O3 = A3 + 1 | A3, G++ = good) &= 1 \\
p(O3 = A3 | A3, G++ = bad) &= 0.98 \\
p(O3 \neq A3 | A3, G++ = bad) &= 0.01 \\
p(O1 = A2 + A3 | A2, A3, G+ = good) &= 1 \\
p(O1 = max(A2, A3) | A2, A3, G+ = bad) &= 0.95 \\
p(O2 \neq max(A2, A3) | A2, A3, G+ = bad) &= 0.01 \\
p(L = red | G = good) &= 0.1 \\
p(L = red | G = bad) &= 0.7
\end{aligned}
$$

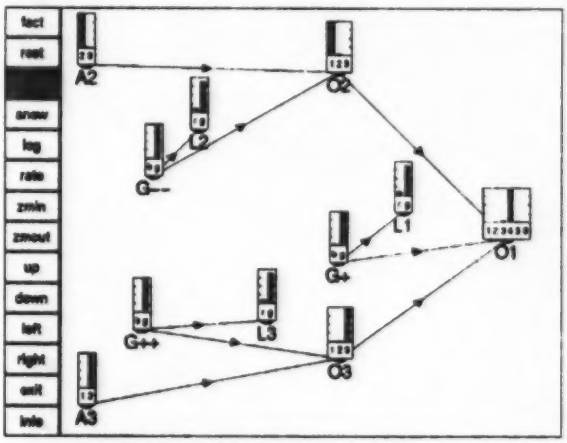

Figure 3: Expected behavior of three gates circuit after inputs A_2, A_3 are specified.

Suppose one knows the inputs are: $A_2 = A_3 = 2$. Figure 3 shows sharp distributions for all gates, which represents one's belief that, without any evidence, everything is expected to be normal. The expected output of the circuit is 4.

Unfortunately, the output is measured as 6. Entering this evidence into the network, PMDs of both G−− and G+ show the possibilities of fault (outcome 'bad' (b) has probability greater than 0.4 in Figure 4). With the understanding of the circuit, one can tell that only 2 cases are quite likely. One is faulty G−− and normal G+ and G++, in which case G−− produces output 3 which explains the error. The other is

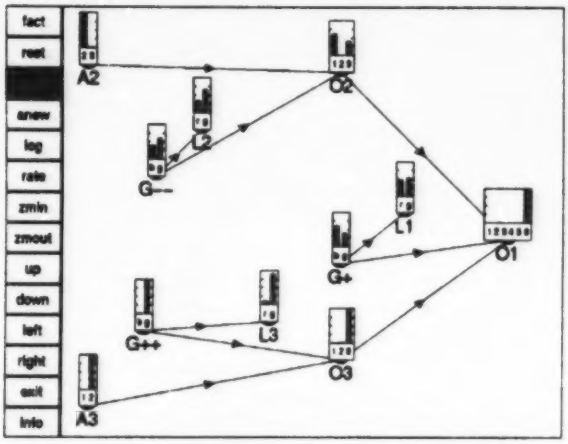

Figure 4: Output $O_1 = 6$ rather than expected 4 is observed.

faulty G+ and normal G−− and G++, in which case G−− produces the expected value 1 and G+ generates the error. Other combinations are quite unlikely. However, without assuming this global knowledge, it is not obvious from the PMDs whether both gates are faulty, or only one of them. Let the lower bound $p_0 = 0.5$. Since none of the outcomes of either gate is greater than 0.75, the threshold by theorem 1, one can not determine the LBCH for the pair of gates. Additional evidence has to be gathered.

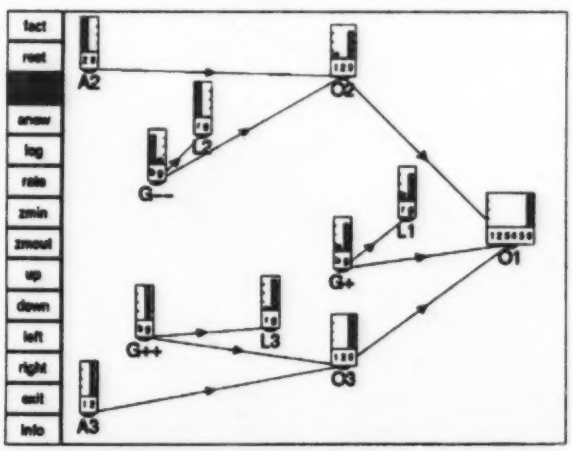

Figure 5: The indicator light L_2 is found to be red (signaling faulty).

After checking the indicator light L_2 and entering the value 'red' (signaling faulty), Figure 5 shows $p(G--=bad|\mathbf{e}) = \mathbf{0.82}$, $p(G+=good|\mathbf{e}) = \mathbf{0.80}$, $p(G++=good|\mathbf{e}) = \mathbf{0.99}$. To determine the LBCH over two variables G−− and G+, this result is sufficient according to the individual threshold of theorem 1. The result is also sufficient for the LBCH over three gate variables according to the sum threshold of theorem 1 (the threshold is 2.5). But it is not sufficient for the LBCH over three gate variables if the individual threshold of theorem 1 is adopted (0.83). Suppose this threshold is used and one gathers another piece of evidence $L_1 = green$ (Figure 6). One finally has $p(G--=bad|\mathbf{e}) = \mathbf{0.93}$, $p(G+=good|\mathbf{e}) = \mathbf{0.92}$, $p(G++=good|\mathbf{e}) = \mathbf{0.99}$, which gives the LBCH over three gate variables by any of the two thresholds. Using auxiliary node method for three gates, the LBCH has probability 0.92.

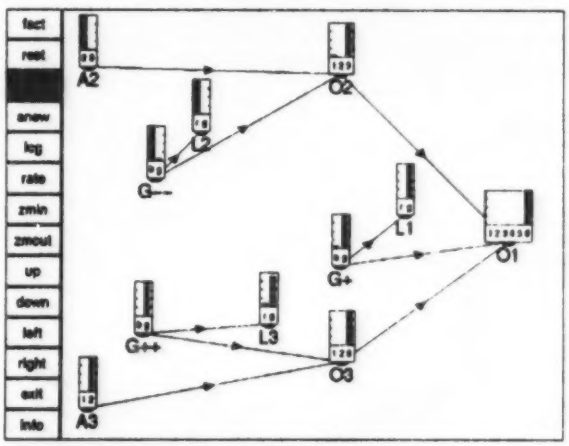

Figure 6: The indicator light L_1 is found to be green (signaling normal).

7 Conclusion

This paper presents a straightforward threshold method of computing the LBCH via belief updating. LBCHs are needed when decisions are to be made where correct decision is associated with reward and wrong decision is associated with penalty. Dominant LBCHs are also MLCHs. Although the threshold method is practical only if the number of variables over which the LBCH is to be computed is small, it

is argued that computing the LBCH over small number of variables can be justified and is often desirable.

The method can be used in conjunction with any belief updating algorithm either based on original networks or based on secondary structures. There has been no known algorithm performing belief revision in clique tree based secondary structures. Such secondary structures have advantages when domain knowledge are acquired once but used for multiple cases (e.g. expert systems); and when the net topology is multiply connected. Therefore, the threshold method provides secondary structure based algorithms with additional capability of obtaining LBCHs. The extra computation required is trivial.

Some features of the method are worth pointing out. First, the method does not provide probability values of LBCHs. In most cases, the actual probabilities of LBCHs are much larger than lower bounds such that high expected utility are associated with the corresponding decisions. Second, before adequate evidence is gathered such that a LBCH can be warranted, no LBCH can be given. When users have the option to collect more evidence, this feature prompts users not to stop with premature judgments. However when users are bounded with limited resource (time or information sources), it is possiole that no LBCH is provided even if the actual MLCH has a probability greater than the lower bound. This is because the method is based on a sufficient condition of LBCHs (Theorem 1) which is not a necessary condition. It seems this is not a major problem since users can always rely on PMDs for rough guidance.

Acknowledgements

This research is supported by Operating Grant OGP0090307 from NSERC, and CRD3474 from the Centre for Systems Science at Simon Fraser University. The author would like to thank David Poole, Nevin Zhange, and Bill Havens for insightful comments to the early manuscript.

References

[Andreassen et al. 89] S. Andreassen, F.V. Jensen, S.K. Andersen, B. Falck, U. Kjerulff, M. Woldbye, A.R. Sorensen, A. Rosenfalck and F. Jensen, "MUNIN - an expert EMG assistant", J.E. Desmedt, (Edt), *Computer-Aided Electromyography and Expert Systems*, Elsevier, 255-277, 1989.

[Heckerman 90] D. Heckerman, A tractable inference algorithm for diagnosing multiple diseases, M. Henrion, R.D. Shachter, L.N. Kanal and J.F. Lemmer. (Eds.), *Uncertainty in Artificial Intelligence 5*, Elsevier Science Publishers, 163-171, 1990.

[Henrion 88] M. Henrion, Propagating uncertainty in Bayesian networks by probabilistic logic sampling, J. F. Lemmer and L. N. Kanal (Eds), *Uncertainty in Artificial Intelligence 2*, Elsevier Science Publishers, 149-163, 1988.

[Henrion 91] M. Henrion, Search-based methods to bound diagnostic probabilities in very large belief nets, B.D. D'Ambrosio, P. Smets and P.P. Bonissone (Eds.), *Proc. 7th Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann, 1991.

[Huchlenbroich 91] P. Huchlenbroich, Methodological principles in medical knowledge programming: part 1, *Artificial Intelligence in Medicine*, Vol.3, No.2, 113-123, 1991.

[Jensen, Lauritzen and Olesen 90a]
F.V. Jensen, S.L. Lauritzen and K.G. Olesen, Bayesian updating in causal probabilistic networks by local computations, *Computational Statistics Quarterly*, 4, 269-282, 1990.

[Jensen, Olesen and Andersen 90b] ,
F.V. Jensen, K.G. Olesen and S.K. Andersen, "An algebra of Bayesian belief universes for knowledge based systems", *Networks*, Vol. 20, 637-659, 1990.

[Lauritzen and Spiegelhalter 88] S.L. Lauritzen and D.J. Spiegelhalter, Local computation with probabilities on graphical structures and their application to expert systems, *Journal of the Royal Statistical Society*, Series B, 50, 157-244, 1988.

[Patil, Szolovits and Schwartz 82] R.S. Patil, P. Szolovits and W.B. Schwartz, "Modeling knowledge of the patient in acid-base and electrolyte disorders", P. Szolovits (Edt), *Artificial Intelligence in Medicine*, 191-226, 1982.

[Pearl 86] J. Pearl, "Fusion, propagation, and structuring in belief networks", *Artificial Intelligence*, 29: 241-288, 1986.

[Pearl 88] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, 1988.

[Poole and Provan 90] D. Poole and G.M. Provan, What is an optimal diagnosis?, *Proc. Sixth Conference on Uncertainty in Artificial Intelligence*, Cambridge, Mass., 46-53, 1990.

[Shachter 86] R.D. Shachter, Evaluating influence diagrams, *Operations Research*, Vol.34, 6, 871-882, 1986.

[Shachter 88] R.D. Shachter, Probabilistic inference and influence diagrams, *Operations Research*, Vol.36, 4, 589-604, 1988.

[von Neumann and Morgenstern 47] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University press, Princeton, NJ, 1947.

[Xiang et al. 91] Y. Xiang, B. Pant, A. Eisen, M.P. Beddoes, and D. Poole, PAINULIM: a neuromuscular diagnostic aid using multiply sectioned Bayesian belief networks, *Proc. ISMM International Conference on Mini and Microcomputers in Medicine and Healthcare*, Long Beach, CA, 64-69, 1991.

[Xiang, Poole and Beddoes 91] Y. Xiang, D. Poole, and M. P. Beddoes, Multiply sectioned Bayesian networks and junction forests for large knowledge based systems, submitted to *Computational Intelligence*, 1991.

Centre for Systems Science

• Simon Fraser University established the **Centre for Systems Science** as a provincial centre for advanced study. The objective of CSS is twofold: coordinate and combine advanced systems research at SFU, and encourage application of research results within industry in order to further the technological development of British Columbia.

• CSS has assembled three research groups to streamline its efforts. The orientation of these groups is as follows: intelligent systems, microelectronics, and computer and communications systems. Because of shared research interests, this latter group merges with LCCR (**Laboratory for Computer and Communications Research**) for the publication of technical reports.

• CSS's three groups involve over 80 skilled faculty researchers principally from SFU's Schools of Engineering Science and Computing Science, but also draw substantially on Mathematics, Physics, Geography, Philosophy, Linguistics, Psychology, Communication, and Kinesiology.